A Reduced-Restricted-Quasi-Newton–Raphson Method for Locating and Optimizing Energy Crossing Points Between Two Potential Energy Surfaces

JOSEP MARIA ANGLADA,[1] JOSEP MARIA BOFILL[2]

[1] *C.I.D.-C.S.I.C., Jordi Girona Salgado 18-26, E-08034 Barcelona, Catalunya, Spain*
[2] *Departament de Química Orgànica, Universitat de Barcelona, Martí i Franquès 1, E-08028 Barcelona, Catalunya, Spain*

Received 7 May 1996; accepted 6 November 1996

ABSTRACT: We present a method for the location and optimization of an intersection energy point between two potential energy surfaces. The procedure directly optimizes the excited state energy using a quasi-Newton–Raphson method coupled with a restricted step algorithm. A linear transformation is also used for the solution of the quasi-Newton–Raphson equations. The efficiency of the algorithm is analyzed and demonstrated in some examples. © 1997 by John Wiley & Sons, Inc. *J Comput Chem* **18**: 992–1003, 1997

Keywords: location of funnels; conical intersections; crossing points; reduced-quasi-Newton–Raphson method; restricted step algorithm

Introduction

In recent years it has been demonstrated that many reactions occur through crossings between two potential energy surfaces. The two surfaces, labeled $S_{f_{ex}}$ and $S_{f_{gs}}$ for the excited and ground state, respectively, are separated by a funnel. Through the funnel the molecule may undergo a radiationless decay from the excited state to the ground state. A funnel corresponds either to a true conical intersection[1] or a weakly avoided crossing. The conical intersection is defined as the region where the surfaces of the states ex and gs, even with the same symmetry, intersect along the $(n - 2)$-dimensional subspace when the energy is plotted against the n nuclear coordinates ($n = 3N -$

Correspondence to: Prof. J. M. Bofill; e-mail: jmbofill@canigo.qo.ub.es.

Contract grant sponsor: DGICYT; contract grant number PB92-0796-C01-02

CCC 0192-8651 / 97 / 080992-12

6).[2] In most nonadiabatic photoreactions the conical intersections are a current feature.

The location of a conical intersection is an equality constraint minimization problem. Practical algorithms exist for the location of crossing points.[3–10] Some of the algorithms fall in the category of the Lagrange–Newton type methods (see refs. 3, 5, 7, 8), while others can be seen as projected gradient methods (e.g., see ref. 9). In the Lagrange–Newton type algorithms, the constraints are introduced by the Lagrangian multipliers; that is, the energy of the state ex, E_{ex}, is minimized subject to two constraints. On the other hand, the projected gradient methods consist of the minimization of E_{ex} using as a gradient the projection of the gradient of E_{ex} in the $(n-2)$-dimensional subspace orthogonal to the 2-dimensional subspace added to a vector that measures the feasibility, namely $(E_{ex} - E_{gs})$. Many of these methods have been successfully applied to quite a large number of problems,[11] but to our knowledge their convergence behavior has only been studied in a very few cases. In general, these methods present an oscillatory behavior that increases the number of iterations[9]; therefore, the location of the conical intersections becomes very expensive.

In this article we present a reduced quasi-Newton–Raphson method for the location and optimization of a conical intersection. We incorporated the restricted step technique to this algorithm,[12] which was successfully applied in other types of optimization problems such as the location of transition structures.[13] At this point it is important to emphasize that both types of problems, finding transition structures (first-order saddle points) and conical intersections, are quite close conceptually and computationally.

First we summarize the theory of the conical intersections, then the mathematical basis of the minimization with constraints is reviewed and an algorithm is described. Finally, some examples are given and the convergence behavior of the algorithm is analyzed from the numerical point of view.

Theoretical Background

THEORY OF CONICAL INTERSECTIONS

This theory has been reviewed several times.[6, 14, 15] Therefore, we will present only a few comments. In a two-level system, the energy levels $\langle \Psi_{gs}|\mathbf{H}|\Psi_{gs}\rangle = E_{gs}$ and $\langle \Psi_{es}|\mathbf{H}|\Psi_{es}\rangle = E_{es}$ are degenerate if $E_{gs} - E_{ex} = 0$ and $\langle \Psi_{gs}|\mathbf{H}|\Psi_{ex}\rangle = H_{gs,ex} = 0$, where $\mathbf{H}$ is the configuration interaction Hamiltonian and Ψ_{gs} and Ψ_{ex} are the corresponding wave functions for the ground state and excited state, respectively. The Taylor expansions of the latter equalities to first order with respect to the nuclear displacement $\Delta\mathbf{q}$ are

$$E_{gs} - E_{ex} = E^0_{gs} - E^0_{ex} + \left(\frac{\partial(E_{gs} - E_{ex})}{\partial \mathbf{q}}\right)^T_{\mathbf{q}=\mathbf{q}_0} \Delta\mathbf{q}_0 = E^0_{gs} - E^0_{ex} + \mathbf{x}_1^T \Delta\mathbf{q}_0 = 0, \quad (1a)$$

$$H_{gs,ex} = H^0_{gs,ex} + \left(\frac{\partial H_{gs,ex}}{\partial \mathbf{q}}\right)^T_{\mathbf{q}=\mathbf{q}_0} \Delta\mathbf{q}_0 = H^0_{gs,ex} + \mathbf{x}_{1,2}^T \Delta\mathbf{q}_0 = 0, \quad (1b)$$

where $\Delta\mathbf{q}_0 = \mathbf{q} - \mathbf{q}_0$. Let us assume that $\mathbf{q}_0$ already satisfies the degeneracy conditions, that is $E^0_{gs} - E^0_{ex} = 0$ and $H^0_{gs,ex} = 0$. Then according to eq. (1) we have

$$\mathbf{x}_1^T \Delta\mathbf{q}_0 = 0, \quad (2a)$$

$$\mathbf{x}_{1,2}^T \Delta\mathbf{q}_0 = 0. \quad (2b)$$

In others words, to preserve the energy degeneracy the variation vector of the nuclear parameters $\Delta\mathbf{q}_0$ should be orthogonal to the subspace spanned by the linear independent vectors $\mathbf{x}_1$ and $\mathbf{x}_{1,2}$.[6, 15, 16] The linear subspace defined by the vectors $\mathbf{x}_1$ and $\mathbf{x}_{1,2}$ will be called branching subspace, S_b, and its orthogonal complement by tangent intersection subspace, S_{ti}, with dimension $n-2$.[15] Clearly the degeneracy will be preserved in the $(n-2)$-dimension S_{ti} subspace and lifted in the 2-dimension S_b subspace. With the previous considerations, to locate a conical intersection of lower energy one must minimize (maximize) the energy E_{ex} (E_{gs}) in the S_{ti} subspace.[6] The generalization of the previous results was given by Katriel and Davidson[16] who considered an m-fold degenerate ground state. In this case eqs. (2) take the following form:

$$\left(\frac{\partial(E_1 - E_i)}{\partial \mathbf{q}}\right)^T_{\mathbf{q}=\mathbf{q}_0} \Delta\mathbf{q}_0 = \mathbf{x}_i^T \Delta\mathbf{q}_0 = 0 \qquad i = 2, \ldots, m; \quad (3a)$$

$$\left(\frac{\partial H_{i,j}}{\partial \mathbf{q}}\right)^T_{\mathbf{q}=\mathbf{q}_0} \Delta\mathbf{q}_0 = \mathbf{x}_{i,j}^T \Delta\mathbf{q}_0 = 0 \qquad i < j = i+1, \ldots, m; i = 1, \ldots, m-1. \quad (3b)$$

The latter equations means that $\Delta\mathbf{q}_0$ has to be

orthogonal to the $(m-1)(m+2)/2$ set of the vectors $\{\mathbf{x}_i\}$ and $\{\mathbf{x}_{i,j}\}$. As pointed out by Katriel and Davidson,[16] for a molecule with n degrees of freedom, the maximum degeneracy is given by the largest m that satisfies the relationship $(m-1)(m+2)/2 \leq n$.

The whole theory is formulated in the quasidiabatic basis but Ragazos et al.[6] showed that in the adiabatic basis one only needs the $\mathbf{x}_1$ vector because in this basis $\mathbf{x}_1$ and $\mathbf{x}_{1,2}$ are related.

MATHEMATICAL BASIS OF THEORY OF CONSTRAINED OPTIMIZATION

The location of conical intersections falls in the set of nonlinear equality constrained optimization problems.[3–10] Essentially the general structure of this type of optimization problems is

$$\underset{q}{\text{minimize}}\, E_{\text{ex}}(\mathbf{q}), \tag{4a}$$

$$\text{subject to} \quad \mathbf{r}(\mathbf{q}) = \mathbf{0}, \tag{4b}$$

where $\mathbf{r}(\mathbf{q})$ is the vector that contains the nonlinear equality constraints. According to the previous discussion, it has the following form for a twofold degenerate ground state:

$$\mathbf{r}(\mathbf{q}) = \begin{pmatrix} E_{\text{gs}} - E_{\text{ex}} \\ H_{\text{gs,ex}} \end{pmatrix}. \tag{5}$$

Very often the set of eqs. (4) are solved using the so-called method of Lagrange multipliers. The method introduces the Lagrangian function,

$$L(\mathbf{q},\mathbf{l}) = E_{\text{ex}}(\mathbf{q}) - \mathbf{l}^T\mathbf{r}(\mathbf{q}), \tag{6}$$

where the $\mathbf{l}$ vectors are the Lagrangian multipliers. The Taylor series expansion to second order for $L(\mathbf{q},\mathbf{l})$ around $\mathbf{q}_0$ and $\mathbf{l}_0$ gives

$$\begin{aligned} L(\mathbf{q}_0 + \Delta\mathbf{q}_0, \mathbf{l}_0 + \Delta\mathbf{l}_0) &= L(\mathbf{q}_0,\mathbf{l}_0) + (\nabla L(\mathbf{q}_0,\mathbf{l}_0))^T \begin{pmatrix} \Delta\mathbf{q}_0 \\ \Delta\mathbf{l}_0 \end{pmatrix} \\ &\quad + \frac{1}{2}(\Delta\mathbf{q}_0, \Delta\mathbf{l}_0)^T \left[\nabla^2 L(\mathbf{q}_0,\mathbf{l}_0)\right] \begin{pmatrix} \Delta\mathbf{q}_0 \\ \Delta\mathbf{l}_0 \end{pmatrix} \\ &= E^0_{\text{ex}} + \Delta\mathbf{q}_0^T\mathbf{g}^0_{\text{ex}} \\ &\quad + \frac{1}{2}\Delta\mathbf{q}_0^T\left(\mathbf{H}^0_{\text{ex}} - \sum_{i=1}^{2} l_{i,0}\nabla^2\mathbf{r}_i(\mathbf{q}_0)\right)\Delta\mathbf{q}_0 \\ &\quad - \mathbf{l}^T\left(\mathbf{r}(\mathbf{q}_0) + [\mathbf{R}(\mathbf{q}_0)]^T\Delta\mathbf{q}_0\right) \\ &= E^0_{\text{ex}} + \Delta\mathbf{q}_0^T\mathbf{g}^0_{\text{ex}} + \frac{1}{2}\Delta\mathbf{q}_0^T\mathbf{W}^0_{\text{ex}}\Delta\mathbf{q}_0 \\ &\quad - \mathbf{l}^T\left(\mathbf{r}(\mathbf{q}_0) + [\mathbf{R}(\mathbf{q}_0)]^T\Delta\mathbf{q}_0\right), \end{aligned} \tag{7}$$

where the vector $\mathbf{l} = \mathbf{l}_0 + \Delta\mathbf{l}_0$, and the vector and matrix $\mathbf{g}^0_{\text{ex}}$ and $\mathbf{H}^0_{\text{ex}}$ are the gradient and Hessian of E_{ex} at $\mathbf{q}_0$, respectively. The $\mathbf{R}(\mathbf{q}_0)$ is the matrix $\mathbf{R}(\mathbf{q}) = [\nabla\mathbf{r}_1(\mathbf{q})\ \nabla\mathbf{r}_2(\mathbf{q})]$ at the point $\mathbf{q}_0$. The $\mathbf{W}^0_{\text{ex}}$ matrix is the Hessian of the Lagrangian function defined in eq. (6). Note that in the present case the $\mathbf{R}(\mathbf{q})$ matrix is

$$\mathbf{R}(\mathbf{q}) = [\mathbf{x}_1\ \mathbf{x}_{1,2}]. \tag{8}$$

If one forgets that in equation (7) the $\mathbf{W}^0_{\text{ex}}$ matrix depends on $\mathbf{l}_0$, then the last equality can be seen as the Lagrangian function, $L(\mathbf{q}_0,\mathbf{l})$, with linear constraint. The latter can be formulated in a more general way as

$$\underset{\Delta\mathbf{q}_0}{\text{minimize}}\, Q_{\text{ex}}(\Delta\mathbf{q}_0) = \Delta\mathbf{q}_0^T\mathbf{g}^0_{\text{ex}} + \frac{1}{2}\Delta\mathbf{q}_0^T\mathbf{W}^0_{\text{ex}}\Delta\mathbf{q}_0, \tag{9a}$$

subject to

$$\mathbf{r}(\mathbf{q}_0) + [\mathbf{R}(\mathbf{q}_0)]^T\Delta\mathbf{q}_0 = \mathbf{0}. \tag{9b}$$

These arguments are the basis of the Han–Powell algorithm[17,18] reviewed by Gabay.[19] In order to solve the set of eqs. (9), we use the generalized elimination method,[12,20] which essentially consists of a linear transformation of the variables. Defining the matrices $\mathbf{T}_{\text{b}}$ and $\mathbf{T}_{\text{ti}}$ of dimension $n \times 2$ and $n \times (n-2)$, respectively, the transformation is given by

$$\Delta\mathbf{q}_0 = [\mathbf{T}_{\text{b}}\ \mathbf{T}_{\text{ti}}]\begin{pmatrix} \Delta\mathbf{y}_0 \\ \Delta\mathbf{x}_0 \end{pmatrix} = \mathbf{T}_{\text{b}}\Delta\mathbf{y}_0 + \mathbf{T}_{\text{ti}}\Delta\mathbf{x}_0, \tag{10}$$

where $\mathbf{x}_0$ and $\mathbf{y}_0$ are the new variables. Note that $\mathbf{x}_0$ and $\mathbf{y}_0$ are the variables associated with the S_{ti} and S_{b} subspaces, respectively. The matrices $\mathbf{T}_{\text{b}}$ and $\mathbf{T}_{\text{ti}}$ have the following properties: $[\mathbf{R}(\mathbf{q})]^T\mathbf{T}_{\text{b}} = \mathbf{I}_{\text{b}}$, $\mathbf{T}_{\text{b}}^T\mathbf{T}_{\text{b}} = \mathbf{I}_{\text{b}}$, $[\mathbf{R}(\mathbf{q})]^T\mathbf{T}_{\text{ti}} = \mathbf{T}_{\text{b}}^T\mathbf{T}_{\text{ti}} = \mathbf{0}_{2\times(n-2)}$ and $\mathbf{T}_{\text{ti}}^T\mathbf{T}_{\text{ti}} = \mathbf{I}_{\text{ti}}$, where $\mathbf{I}_{\text{b}}$ is the unit matrix of the S_{b} subspace and $\mathbf{I}_{\text{ti}}$ is the unit matrix of the S_{ti} subspace. The $\mathbf{0}_{2\times(n-2)}$ matrix is a zero matrix of $2 \times (n-2)$ dimension. Through the above relationships both the $\mathbf{T}_{\text{b}}$ and $\mathbf{T}_{\text{ti}}$ matrices depend on $\mathbf{q}$; consequently at $\mathbf{q}_0$ they will be denoted as $\mathbf{T}^0_{\text{b}}$ and $\mathbf{T}^0_{\text{ti}}$, respectively. On the other hand, if the set of vectors $\{\nabla\mathbf{r}_i(\mathbf{q})\}$ are linearly independent, then

the matrix $[\mathbf{T}_b \quad \mathbf{T}_{ti}]$ is nonsingular.[20] We emphasize that for this type of problem a point $\mathbf{q}^*$ is the solution of eqs. (4) if it satisfies the following necessary and sufficient conditions:

1. The restriction should be satisfied at $\mathbf{q}^*$, $\mathbf{r}(\mathbf{q}^*) = \mathbf{0}$.
2. The point $\mathbf{q}^*$ is stationary in the S_{ti} subspace, that is $(\mathbf{T}_{ti}^*)^T \mathbf{g}_{ex}(\mathbf{q}^*) = \mathbf{0}$, where $\mathbf{T}_{ti}^*$ is the matrix $\mathbf{T}_{ti}$ at $\mathbf{q}^*$.
3. The stationary point $\mathbf{q}^*$ has a character of minimum if the matrix $(\mathbf{T}_{ti}^*)^T \mathbf{W}_{ex}^* \mathbf{T}_{ti}^*$ is positive definite, where $\mathbf{W}_{ex}^*$ is the matrix $\mathbf{W}_{ex}$ computed at $\mathbf{q}^*$.

Now first substituting eq. (10) in eq. (9b) we get

$$\Delta \mathbf{y}_0 = -\mathbf{r}(\mathbf{q}_0), \tag{11}$$

and again substituting eq. (10) in eq. (9a) and taking into account eq. (11), we obtain the following unconstrained quadratic optimization problem:

$$\begin{aligned} \underset{\Delta \mathbf{x}_0}{\text{minimize}}\, & Q_{ex}(\Delta \mathbf{x}_0) \\ &= (\mathbf{T}_{ti}^0 \Delta \mathbf{x}_0 - \mathbf{T}_b^0 \mathbf{r}(\mathbf{q}_0))^T \mathbf{g}_{ex}^0 \\ &\quad - (\mathbf{r}(\mathbf{q}_0))^T (\mathbf{T}_b^0)^T \mathbf{W}_{ex}^0 \mathbf{T}_{ti}^0 \Delta \mathbf{x}_0 \\ &\quad + \frac{1}{2} (\mathbf{r}(\mathbf{q}_0))^T (\mathbf{T}_b^0)^T \mathbf{W}_{ex}^0 \mathbf{T}_b^0 \mathbf{r}(\mathbf{q}_0) \\ &\quad + \frac{1}{2} \Delta \mathbf{x}_0^T (\mathbf{T}_{ti}^0)^T \mathbf{W}_{ex}^0 \mathbf{T}_{ti}^0 \Delta \mathbf{x}_0. \end{aligned} \tag{12}$$

Solving eq. (12) with respect to $\Delta \mathbf{x}_0$ and substituting the result in eq. (10) we obtain

$$\begin{aligned} \Delta \mathbf{q}_0 = & -\mathbf{T}_{ti}^0 \left[(\mathbf{T}_{ti}^0)^T \mathbf{W}_{ex}^0 \mathbf{T}_{ti}^0 \right]^{-1} \Big((\mathbf{T}_{ti}^0)^T \mathbf{g}_{ex}^0 \\ & - (\mathbf{T}_{ti}^0)^T \mathbf{W}_{ex}^0 \mathbf{T}_b^0 \mathbf{r}(\mathbf{q}_0) \Big) - \mathbf{T}_b^0 \mathbf{r}(\mathbf{q}_0). \end{aligned} \tag{13}$$

Equation (13) is the formal solution of the set of eqs. (9). The matrix $[(\mathbf{T}_{ti})^T \mathbf{W}_{ex} \mathbf{T}_{ti}]$ is the so-called reduced Hessian matrix and $(\mathbf{T}_{ti})^T (\mathbf{g}_{ex} - \mathbf{W}_{ex} \mathbf{T}_b \mathbf{r}(\mathbf{q}))$ is the reduced gradient, which takes into account the feasibility of the restriction and the energy minimization on the S_{ti} subspace. The solution of eq. (12) is found through a quasi-Newton method.[12] Consequently, rather than using the $\mathbf{W}_{ex}^0$ matrix, one uses an approximation to it represented by the $\mathbf{B}_{ex}^0$ matrix. The $\mathbf{B}_{ex}$ matrix is updated at each iteration using the quasi-Newton condition[12] applied to the Lagrangian function given in eq. (6), which is

$$\begin{aligned} \nabla L(\mathbf{q}_0 + \Delta \mathbf{q}_0, \mathbf{l}_0 + \Delta \mathbf{l}_0) &= \nabla L(\mathbf{q}_0, \mathbf{l}_0) \\ &\quad + \left[\nabla^2 L(\mathbf{q}_0, \mathbf{l}_0) + \mathbf{E}(\mathbf{q}_0, \mathbf{l}_0) \right] \begin{pmatrix} \Delta \mathbf{q}_0 \\ \Delta \mathbf{l}_0 \end{pmatrix}, \end{aligned} \tag{14}$$

where $\mathbf{E}(\mathbf{q}_0, \mathbf{l}_0)$ is a matrix correction to be determined and it takes into account the error due to the truncation until first order of the Taylor series expansion of $\nabla L(\mathbf{q}_0, \mathbf{l}_0)$ vector.[12] In matrix form, expression (14) is

$$\begin{aligned} \begin{pmatrix} \mathbf{g}_{ex} - \mathbf{R}(\mathbf{q})\mathbf{l} \\ -\mathbf{r}(\mathbf{q}) \end{pmatrix} = & \begin{pmatrix} \mathbf{g}_{ex}^0 - \mathbf{R}(\mathbf{q}_0)\mathbf{l}_0 \\ -\mathbf{r}(\mathbf{q}_0) \end{pmatrix} \\ & + \left[\begin{pmatrix} \mathbf{B}_{ex}^0 & -\mathbf{R}(\mathbf{q}_0) \\ -(\mathbf{R}(\mathbf{q}_0))^T & \mathbf{0} \end{pmatrix} \right. \\ & \left. + \begin{pmatrix} \mathbf{E}_{11}^0 & \mathbf{0} \\ \mathbf{0}^T & \mathbf{0} \end{pmatrix} \right] \begin{pmatrix} \Delta \mathbf{q}_0 \\ \Delta \mathbf{l}_0 \end{pmatrix}, \end{aligned} \tag{15}$$

and taking into account the first row of eq. (15) we get

$$\begin{aligned} \mathbf{g}_{ex} - \mathbf{g}_{ex}^0 - (\mathbf{R}(\mathbf{q}) - \mathbf{R}(\mathbf{q}_0))\mathbf{l} &= \mathbf{h}_{ex} - \mathbf{h}_{ex}^0 = \left[\mathbf{B}_{ex}^0 + \mathbf{E}_{11}^0 \right] \Delta \mathbf{q}_0 = \mathbf{B}_{ex} \Delta \mathbf{q}_0, \end{aligned} \tag{16}$$

where $\mathbf{h}_{ex}$ and $\mathbf{h}_{ex}^0$ are the gradients of the Lagrangian function (6) with respect to $\mathbf{q}$ at $(\mathbf{q}, \mathbf{l})$ and $(\mathbf{q}_0, \mathbf{l})$, respectively, and the gradient vector $\mathbf{g}_{ex} = \mathbf{g}_{ex}(\mathbf{q}) = \mathbf{g}_{ex}(\mathbf{q}_0 + \Delta \mathbf{q}_0)$. Note that the $\mathbf{E}_{11}^0$ matrix is the nonzero part of the $\mathbf{E}(\mathbf{q}_0, \mathbf{l}_0)$ matrix correction. The $\mathbf{l}$ vector of the Lagrangian multipliers is computed as

$$\mathbf{l} = (\mathbf{T}_b^0)^T \mathbf{g}_{ex}, \tag{17}$$

which is merely a first-order estimation of the Lagrangian multiplier vector $\mathbf{l}^*$ at the solution. In fact the vector given by eq. (17) contains the Lagrangian multipliers at the solution of the quadratic problem (9) that defines $\Delta \mathbf{q}_0$. The $\mathbf{E}_{11}^0$ matrix is evaluated in the usual way by the variable metric methods.[12] Depending on the evaluation of the $\mathbf{E}_{11}^0$ matrix, one gets the Broyden–Fletcher–Goldfarb–Shanno (BFGS)[12] or the Murtagh–Sargent–Powell (MSP)[21] formula for the correction of

the $\mathbf{B}_{ex}$ matrix. These formulae are

$$\mathbf{B}_{ex} = \mathbf{B}_{ex}^0 + \mathbf{E}_{11}^0 = \mathbf{B}_{ex}^0 + \frac{(\mathbf{h}_{ex} - \mathbf{h}_{ex}^0)(\mathbf{h}_{ex} - \mathbf{h}_{ex}^0)^T}{(\mathbf{h}_{ex} - \mathbf{h}_{ex}^0)^T \Delta\mathbf{q}_0} - \frac{\mathbf{B}_{ex}^0 \Delta\mathbf{q}_0 (\Delta\mathbf{q}_0)^T \mathbf{B}_{ex}^0}{(\Delta\mathbf{q}_0)^T \mathbf{B}_{ex}^0 \Delta\mathbf{q}_0} \quad (18)$$

for the BFGS[12] and

$$\mathbf{B}_{ex} = \mathbf{B}_{ex}^0 + \mathbf{E}_{11}^0 = \mathbf{B}_{ex}^0 + (1 - \phi)\frac{\mathbf{j}_0 \mathbf{j}_0^T}{(\Delta\mathbf{q}_0)^T \mathbf{j}_0} + \phi\left[\frac{\Delta\mathbf{q}_0 \mathbf{j}_0^T + \mathbf{j}_0 (\Delta\mathbf{q}_0)^T}{(\Delta\mathbf{q}_0)^T \Delta\mathbf{q}_0} - \frac{(\Delta\mathbf{q}_0)^T \mathbf{j}_0}{\left((\Delta\mathbf{q}_0)^T \Delta\mathbf{q}_0\right)^2} \Delta\mathbf{q}_0 (\Delta\mathbf{q}_0)^T\right] \quad (19)$$

for the MSP[21] correction, where ϕ is a scalar such that $0 \le \phi \le 1$ and

$$\mathbf{j}_0 = (\mathbf{h}_{ex} - \mathbf{h}_{ex}^0) - \mathbf{B}_{ex}^0 \Delta\mathbf{q}_0. \quad (20)$$

DESCRIPTION AND DETAILS OF ALGORITHM

The mathematical theory presented above is the basis of the following algorithm where the energy of the excited state, E_{ex}, is optimized using eq. (13), changing $\mathbf{W}_{ex}^0$ by $\mathbf{B}_{ex}^0$. Its generalization to an m-fold degenerate ground state problem is trivial.

Given a $\mathbf{q}_0$, compute $E_{ex}(\mathbf{q}_0)$, $E_{gs}(\mathbf{q}_0)$, $\mathbf{g}_{gs}^0$, $\mathbf{g}_{ex}^0$. Set $k = 0$.

1. Compute the $\mathbf{x}_1$ and $\mathbf{x}_{1,2}$ vectors. Construct $\mathbf{r}(\mathbf{q}_k)$ according to eq. (5). Using the Gram–Schmidt process, orthonormalize the set of linearly independent vectors $[\mathbf{x}_1\ \mathbf{x}_{1,2}\ \mathbf{u}_3 \cdots \mathbf{u}_n]$, where the $\mathbf{u}_i^T = (0, \ldots, 1, \ldots, 0)$; the 1 is in the i position. The first two orthonormalized vectors define the $\mathbf{T}_b^k$ matrix and the rest of the $n - 2$ orthonormalized vectors the $\mathbf{T}_{ti}^k$ matrix.
2. Make the transformations $(\mathbf{T}_{ti}^k)^T \mathbf{B}_{ex}^k \mathbf{T}_{ti}^k$ and $(\mathbf{T}_{ti}^k)^T \mathbf{B}_{ex}^k \mathbf{T}_b^k$. Evaluate the reduced gradient in the S_{ti} subspace $\mathbf{g}_{ti}^k = (\mathbf{T}_{ti}^k)^T \mathbf{g}_{ex}^k$. Compute $\Delta\mathbf{q}_k$ according to eq. (13).
3. If $\|(\Delta\mathbf{q}_k)^T \Delta\mathbf{q}_k - (\mathbf{r}(\mathbf{q}_k))^T \mathbf{r}(\mathbf{q}_k)\| = \|(\Delta\mathbf{x}_k)^T \Delta\mathbf{x}_k\| > R_{ti}^k$ or $(\mathbf{T}_{ti}^k)^T \mathbf{B}_{ex}^k \mathbf{T}_{ti}^k$ is not positive definite, then solve the following set of equations on ν_k and $\Delta\mathbf{x}_k$:

$$\left[(\mathbf{T}_{ti}^k)^T \mathbf{B}_{ex}^k \mathbf{T}_{ti}^k + \nu_k \mathbf{I}_{ti}\right]\Delta\mathbf{x}_k = -\left((\mathbf{T}_{ti}^k)^T \mathbf{g}_{ex}^k - (\mathbf{T}_{ti}^k)^T \mathbf{B}_{ex}^k \mathbf{T}_b^k \mathbf{r}(\mathbf{q}_k)\right), \quad (21a)$$

$$(\Delta\mathbf{x}_k)^T \Delta\mathbf{x}_k = (R_{ti}^k)^2. \quad (21b)$$

Using the new $\Delta\mathbf{x}_k$ and eqs. (10) and (11), compute the corrected $\Delta\mathbf{q}_k$.
4. Compute $E_{ex}(\mathbf{q}_k + \Delta\mathbf{q}_k)$ and $\mathbf{r}(\mathbf{q}_k + \Delta\mathbf{q}_k)$ and the ratio

$$r_k = \frac{E_{ex}(\mathbf{q}_k + \Delta\mathbf{q}_k) - E_{ex}(\mathbf{q}_k)}{Q_{ex}(\Delta\mathbf{q}_k)}; \quad (22)$$

if $r_k < r_1$, set $R_{ti}^{k+1} = R_{ti}^k / S_f$;

if $r_k > r_u$ and $\|(\Delta\mathbf{x}_k)^T \Delta\mathbf{x}_k\| = R_{ti}^k$ and

$$\|(\mathbf{r}(\mathbf{q}_k + 1))^T \mathbf{r}(\mathbf{q}_k + 1)\| \le \|(\mathbf{r}(\mathbf{q}_k))^T \mathbf{r}(\mathbf{q}_k)\|,$$

set $R_{ti}^{k+1} = R_{ti}^k \cdot (S_f)^{1/2}$;

otherwise set $R_{ti}^{k+1} = R_{ti}^k$;

if $R_{ti}^{k+1} > R_{max}^{ti}$, set $R_{ti}^{k+1} = R_{max}^{ti}$;

if $r_k \le L_b$, set $\mathbf{q}_{k+1} = \mathbf{q}_k$, $\mathbf{T}_{ti}^{k+1} = \mathbf{T}_{ti}^k$,

$$\mathbf{T}_b^{k+1} = \mathbf{T}_b^k,\quad \mathbf{g}_{ex}^{k+1} = \mathbf{g}_{ex}^k,\quad \mathbf{B}_{ex}^{k+1} = \mathbf{B}_{ex}^k,$$

$$E_{ex}(\mathbf{q}_{k+1}) = E_{ex}(\mathbf{q}_k), \text{ and } k = k + 1;$$

go to 3.
5. Check the convergence on the root mean square (RMS) of $\|(\mathbf{g}_{ti}^k)^T \mathbf{g}_{ti}^k/(n-2)\|$ and $\|(\mathbf{r}(\mathbf{q}_k))^T \mathbf{r}(\mathbf{q}_k)/2\|$; if it is fulfilled stop.
6. Set $\mathbf{q}_{k+1} = \mathbf{q}_k + \Delta\mathbf{q}_k$. Compute $E_{ex}(\mathbf{q}_{k+1})$, $E_{gs}(\mathbf{q}_{k+1})$, $\mathbf{g}_{ex}(\mathbf{q}_{k+1})$, $\mathbf{g}_{gs}(\mathbf{q}_{k+1})$. Update $\mathbf{B}_{ex}^{k+1}$ using either eq. (18) or (19) and the Lagrangian multipliers by eq. (17); that is, $\mathbf{l}_k = (\mathbf{T}_b^k)^T \mathbf{g}_{ex}(\mathbf{q}_{k+1})$, set $k = k + 1$, and go to 1.

The expression $\|\cdot\|$ denotes the Euclidean norm. The parameters r_1, r_u, R_{max}^{ti}, L_b, and S_f are arbitrary and the algorithm is insensible to their change. Suggested values are $r_1 = 0.25$, $r_u = 0.75$, $R_{max}^{ti} = 0.5$, $L_b = 0$, and $S_f = 2$. The above algo-

rithm was formulated for a general diabatic wave function, but if one uses adiabatic wave functions for the case of $m = 2$, then only the $\mathbf{x}_1$ vector is needed and the $\mathbf{r}(\mathbf{q})$ vector contains only an element, $E_{gs} - E_{ex}$.

Now we comment on some parts of the algorithm. The Gram–Schmidt process used in step 1 orthonormalizes the directions of the full space $\{\mathbf{x}_1\ \mathbf{x}_{1,2}\ \mathbf{u}_3 \cdots \mathbf{u}_n\}$, which are not collinear with the eigenvectors of the Hessian matrix $\mathbf{W}_{ex}^0$. This fact is the main difference with the problem of finding transition structures, because in this case the transition vector that characterizes a transition structure is an eigenvector of the Hessian matrix. As will be seen, the use of the orthonormalized directions has some advantages.

The justification of steps 3 and 4 is the following: if within a given threshold the restriction is almost satisfied, then the algorithm is only concerned with the optimization of the energy. In this situation it is important to preserve the restriction during the rest of the optimization. If $\|(\Delta\mathbf{q}_0)^T\Delta\mathbf{q}_0\|$ is big, then it is possible to again lose the restriction causing a deterioration of the process. To avoid this difficulty one should optimize expression (12) until $(\Delta\mathbf{x}_0)^T(\Delta\mathbf{x}_0) \le (\Delta\mathbf{q}_0)^T\Delta\mathbf{q}_0 - \mathbf{r}(\mathbf{q}_0)^T\mathbf{r}(\mathbf{q}_0) = R_0^2 - \mathbf{r}(\mathbf{q}_0)^T\mathbf{r}(\mathbf{q}_0) = (R_{ti}^0)^2$; that is, solve the Lagrangian function $L(Q_{ex}(\Delta\mathbf{x}_0), \nu_0) = Q_{ex}(\Delta\mathbf{x}_0) + \nu_0/2[(\Delta\mathbf{x}_0)^T(\Delta\mathbf{x}_0) - (R_{ti}^0)^2]$. The solution of this problem is given by the set of eqs. (21).[12] This type of restriction step can be handled in this way due to the orthonormalization of the basis $[\mathbf{T}_b^0\ \mathbf{T}_{ti}^0]$. On the other hand, the radius R_{ti}^0 is changed accordingly to the value of $\mathbf{r}(\mathbf{q}_0)^T\mathbf{r}(\mathbf{q}_0)$ and the ratio given by eq. (22). The use of restricted step is justified because in these situations $\mathbf{r}(\mathbf{q}_0 + \Delta\mathbf{q}_0) \approx \mathbf{r}(\mathbf{q}_0) + [\mathbf{R}(\mathbf{q}_0)]^T\Delta\mathbf{q}_0$, which means that the linearization of the constraint is a very good approximation. Then taking into account either eqs. (9a) or (12) we can write

$$E_{ex}(\mathbf{q}_0 + \Delta\mathbf{q}_0) - E_{ex}(\mathbf{q}_0) \approx Q_{ex}(\Delta\mathbf{q}_0) = Q_{ex}(\Delta\mathbf{x}_0), \quad (23)$$

which is the basis of eq. (22) and the restrictions imposed in step 4 of the algorithm.

Regarding the updating of the Hessian matrix, $\mathbf{B}_{ex}$, we suggest using the BFGS[12] formula rather than MSP one,[21] because we are searching a minimum and for these situations BFGS provides better results than the MSP formula.[12, 21] However, there is no guarantee in this algorithm that $(\mathbf{h}_{ex} - \mathbf{h}_{ex}^0)^T\Delta\mathbf{q}_0 > 0$, which is a necessary and sufficient condition using eq. (18) for $\mathbf{B}_{ex}$ to be positive definite if $\mathbf{B}_{ex}^0$ is.[12] This is due to the curvature of the restrictions. Powell[18] suggests a device that consists of replacing $(\mathbf{h}_{ex} - \mathbf{h}_{ex}^0)$ in eq. (18) by the vector

$$\mathbf{z}_{ex}^0 = \theta_0(\mathbf{h}_{ex} - \mathbf{h}_{ex}^0) + (1 - \theta_0)\mathbf{B}_{ex}^0\Delta\mathbf{q}_0, \quad (24)$$

where θ_0 is a scalar between 0 and 1 chosen according to

$$\theta_0 = \begin{cases} 1, & \text{if } (\mathbf{h}_{ex} - \mathbf{h}_{ex}^0)^T\Delta\mathbf{q}_0 \ge \sigma\Delta\mathbf{q}_0^T\mathbf{B}_{ex}^0\Delta\mathbf{q}_0; \\ \dfrac{(1-\sigma)\Delta\mathbf{q}_0^T\mathbf{B}_{ex}^0\Delta\mathbf{q}_0}{\Delta\mathbf{q}_0^T\mathbf{B}_{ex}^0\Delta\mathbf{q}_0 - (\mathbf{h}_{ex} - \mathbf{h}_{ex}^0)^T\Delta\mathbf{q}_0}, & \text{if } (\mathbf{h}_{ex} - \mathbf{h}_{ex}^0)^T\Delta\mathbf{q}_0 < \sigma\Delta\mathbf{q}_0^T\mathbf{B}_{ex}^0\Delta\mathbf{q}_0, \end{cases} \quad (25)$$

with $0 \le \sigma \le 0.5$. Normally $\sigma = 0.2$.[18] Then $(\mathbf{z}_{ex}^0)^T\Delta\mathbf{q}_0 > 0$; hence eq. (18) preserves positive definiteness.

Examples, Analysis, and Discussion

The above algorithm was implemented in the semiempirical program package AMPAC.[22] The following calculations were carried out with the AM1 Hamiltonian.[23] The Hessian matrices at the starting geometries were computed by finite differences of analytic gradients[24] of the energy of the excited state, E_{ex}. The convergence criteria were taken on the RMS gradient of the S_{ti} subspace, $\|(\mathbf{g}_{ti})^T\mathbf{g}_{ti}/(n-1)\|$, as well as on the constraints $\|\mathbf{r}(\mathbf{q})^T\mathbf{r}(\mathbf{q})/1\|$, with the values $8.4 \cdot 10^{-5}$ hartrees/bohr and $6.4 \cdot 10^{-5}$ hartrees, respectively.

The examples presented involve crossing points between singlet–singlet, triplet–triplet, and singlet–triplet electronic states. The appropriate configuration interaction wave function, denoted by

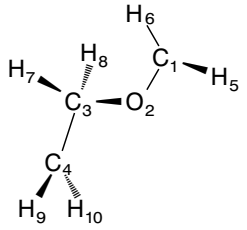
CI (number of electrons, number of molecular orbitals)/AM1, was taken in each case. This wave function consists of a full CI in the subspace defined by the selected orbitals and electrons. This subspace is known as active space and its orbitals as active orbitals. The orbitals used to build the CI wave function are the Hartree–Fock (HF) type orbitals. The active orbitals are mainly selected by taking into account the orbitals implicated in the bond breaking and bond formation associated with the process under study. The analytical gradients are computed by solving the so-called coupled perturbed HF equations (CPHF).[24]

EXAMPLE 1: SINGLET–SINGLET TRANS CONICAL INTERSECTION FOR CARBON–OXYGEN ATTACK OF PATERNO–BUCHI REACTION

This crossing point was reported recently by Palmer et al.[25] in their *ab initio* study of the singlet and triplet Paterno–Buchi reaction with the model system formaldehyde plus ethylene. The calculations were carried out with a CI (6, 5)/AM1 wave function. The starting and final geometries are given in Table I. The behavior of the method is shown in Table II. The BFGS formula was used for the revision of the Hessian matrix at each iteration. The final converged gradient difference vector, $\partial(E_{gs} - E_{ex})/\partial \mathbf{q}$, corresponds essentially to an increase of the O_2C_3 bond distance and a decrease of the bond distances C_1O_2 and C_3C_4. At the final converged point the maximum gradient component in the S_{ti} subspace is $\max|(\mathbf{g}_{ti})_i| = 4.00 \cdot 10^{-6}$ hartrees/bohr. The geometry differs very little from that reported by Palmer et al.,[25] except in the bond distance O_2C_3, which is 0.2 Å smaller.

During the optimization process (see Table II) the restricted step given by the set of eqs. (21) is active in the first three iterations. This behavior is normal in any optimization procedure because in the first iterations the geometry is far from the optimum one and the process has to be controlled. On the other hand, it is worth noting that the RMS gradient criteria is fulfilled much faster than the restriction. A possible reason for this is that the

TABLE I.
Geometry of Singlet – Singlet Trans Conical Intersection for Carbon – Oxygen Attack of Paterno – Buchi Reaction within C_S Symmetry.



Parameter	Initial	Final
C_1O_2	1.278	1.291
O_2C_3	1.400	1.512
C_3C_4	1.428	1.441
H_5C_1	1.077	1.061
H_6C_1	1.091	1.082
H_7C_3	1.125	1.115
H_9C_4	1.089	1.089
$C_1O_2C_3$	137.3	126.1
$O_2C_3C_4$	106.2	103.1
$H_5C_1O_2$	119.0	118.7
$H_6C_1O_2$	121.8	120.9
$H_7C_3C_4$	113.3	114.3
$H_9C_4C_3$	120.4	120.4
$C_1O_2C_3C_4$	180.0	180.0
$H_5C_1O_2C_3$	180.0	180.0
$H_6C_1O_2C_3$	0.0	0.0
$H_7C_3C_4O_2$	118.4	115.8
$H_9C_4C_3O_2$	−83.8	−85.8

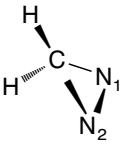
Distances in angstroms and angles in degrees.

TABLE II.
Behavior of Optimization Process for Singlet – Singlet Trans Conical Intersection for Carbon – Oxygen Attack of Paterno – Buchi Reaction.

Iteration	RMS Gradient[a]	$\lvert E_{gs} - E_{ex}\rvert$[b]
1[c]	$1.96 \cdot 10^{-2}$	$2.07 \cdot 10^{-4}$
2[c]	$1.26 \cdot 10^{-2}$	$9.60 \cdot 10^{-5}$
3[c]	$5.49 \cdot 10^{-3}$	$1.60 \cdot 10^{-5}$
4	$3.47 \cdot 10^{-3}$	$1.43 \cdot 10^{-4}$
5	$3.82 \cdot 10^{-3}$	$1.43 \cdot 10^{-4}$
6	$4.05 \cdot 10^{-3}$	$1.28 \cdot 10^{-4}$
7	$4.20 \cdot 10^{-3}$	$1.28 \cdot 10^{-4}$
8	$4.04 \cdot 10^{-3}$	$1.12 \cdot 10^{-4}$
9	$2.98 \cdot 10^{-3}$	$1.12 \cdot 10^{-4}$
10	$1.41 \cdot 10^{-3}$	$9.60 \cdot 10^{-5}$
11	$3.96 \cdot 10^{-4}$	$9.60 \cdot 10^{-5}$
12	$7.60 \cdot 10^{-5}$	$9.60 \cdot 10^{-5}$
13	$1.70 \cdot 10^{-5}$	$9.60 \cdot 10^{-5}$
14	$3.00 \cdot 10^{-6}$	$8.00 \cdot 10^{-5}$
15	$6.00 \cdot 10^{-6}$	$8.00 \cdot 10^{-5}$
16	$4.00 \cdot 10^{-6}$	$8.00 \cdot 10^{-5}$
17	$2.00 \cdot 10^{-6}$	$8.00 \cdot 10^{-5}$
18	$1.00 \cdot 10^{-6}$	$6.4 \cdot 10^{-5}$
19	$2.00 \cdot 10^{-6}$	$6.4 \cdot 10^{-5}$

[a] RMS gradient in hartrees / bohr.
[b] Energy differences in hartrees.
[c] Iterations where the restricted step given by eqs. (21a, b) is active.

TABLE III.
Geometry of Triplet – Triplet Conical Intersection, (^{3}n − π*/^{3}n − σ*), of Diazirine within C_S Symmetry.



Parameter	Final
CN_1	1.415
N_1N_2	1.237
HC	1.108
CN_1N_2	80.4
HCN_1	120.3
HCN_1N_2	98.0

Distances in angstroms and angles in degrees.

algorithm gives quasilinear treatment of the constraint [see eq. (9b)]. Finally we observe that in this example $(\mathbf{h}_{ex} - \mathbf{h}^0_{ex})^T \Delta\mathbf{q}_0 > 0$ remains positive definite during all the process, so the Powell device, eq. (25), was not used.

EXAMPLE 2: TRIPLET–TRIPLET CONICAL INTERSECTION, (^{3}n − π*/^{3}n − σ*), OF DIAZIRINE

This conical intersection was first studied at the *ab initio* level by Bigot et al.[26] and more recently by Yamamoto et al.,[27] who reported a C_1 symmetry geometry for the conical intersection. The calculations were carried out with a CI (4, 4)/AM1 wave function. In Tables III and IV we show the final geometry and the behavior of the method, respectively. The initial geometry was taken from the work of Yamamoto et al.[27] This geometry was first optimized with a fixed CN_1N_2 angle for the second triplet state. The final geometry converged to a C_S symmetry one, which is a point of the seam between the triplet $^3n - \pi^*$ and $^3n - \sigma^*$ electronic surfaces (see the value of $|E_{gs} - E_{ex}|$ in Table IV for the first iteration). Giving this point to the present algorithm it converged within six iterations (see Table IV). Note that the restricted step was inactive during the process. The final converged gradient difference vector, $\partial(E_{gs} - E_{ex})/\partial\mathbf{q}$, corresponds essentially to an increase of the CN_1N_2 bond angle. At the final converged point the maximum gradient component in the S_{ti} subspace is $\max|(\mathbf{g}_{ti})_i| = 1.25 \cdot 10^{-4}$ hartrees/bohr. The final geometry differs very little from that reported by Yamamoto et al.[27] except for a bond distance CN_1, which is about 0.1 Å smaller.

EXAMPLE 3: TRIPLET–TRIPLET TRANS CONICAL INTERSECTION FOR CARBON–CARBON ATTACK OF PATERNO–BUCHI REACTION

In Table V we report the starting and final geometry for this conical intersection. A CI (4, 4)/AM1 wave function was used. The behavior of the method is shown in Table VI. As in the previous examples, the BFGS formula was used for the revision of the Hessian matrix. The main components of the final converged gradient vector, $\partial(E_{gs} - E_{ex})/\partial\mathbf{q}$, corresponds to a decrease of the O_1C_2 bond distance, an increase of the C_2C_3 bond distance, and a decrease of the $O_1C_2C_3$ bond angle. At the final converged point the maximum gradient component in the S_{ti} subspace is $\max|(\mathbf{g}_{ti})_i| = 2.60 \cdot 10^{-5}$ hartrees/bohr.

As in the case of the singlet–singlet trans conical intersection presented above, the optimization process used the restricted step in the first five iterations. Again the Powell device was never used. Finally, we again observe that the RMS gradient criteria is satisfied much faster that the restriction. This is the main reason for the large number of iterations used to achieve the final convergence.

EXAMPLE 4: SINGLET–TRIPLET, $S_0 - T_1$, INTERSECTION IN CARBON–CARBON BOND-BREAKING REGION FOR 1,2-DIOXETANE DECOMPOSITION

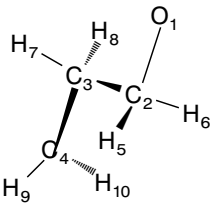
The chemiluminescent decomposition of 1,2-dioxetanes was studied by Reguero et al.[28] at the *ab initio* level. According to this study, the mechanism involves the following steps: a ring opening

TABLE IV.
Behavior of Optimization Process for Triplet – Triplet Conical Intersection, (^{3}n − π*/^{3}n − σ*), of Diazirine.

Iteration	RMS Gradient[a]	$\lvert E_{gs} - E_{ex}\rvert$[b]
1	$2.02 \cdot 10^{-4}$	$8.00 \cdot 10^{-5}$
2	$1.60 \cdot 10^{-4}$	$4.80 \cdot 10^{-5}$
3	$1.43 \cdot 10^{-4}$	$4.80 \cdot 10^{-5}$
4	$1.26 \cdot 10^{-4}$	$4.80 \cdot 10^{-5}$
5	$1.01 \cdot 10^{-4}$	$4.80 \cdot 10^{-5}$
6	$8.40 \cdot 10^{-5}$	$4.80 \cdot 10^{-5}$

[a] RMS gradient in hartrees/bohr.
[a] Energy differences in hartrees.

TABLE V.
Geometry of Triplet – Triplet Trans Conical Intersection for Carbon – Carbon Attack of Paterno – Buchi Reaction within C_S Symmetry.



Parameter	Initial	Final
O_1C_2	1.348	1.362
C_2C_3	1.560	1.514
C_3C_4	1.463	1.472
H_5C_2	1.130	1.129
H_7C_3	1.120	1.123
H_9C_4	1.088	1.085
$O_1C_2C_3$	112.7	108.7
$C_2C_3C_4$	109.9	109.4
$H_5C_2O_1$	108.1	108.5
$H_7C_3C_4$	111.2	110.0
$H_9C_4C_3$	120.3	120.4
$O_1C_2C_3C_4$	180.0	180.0
$H_5C_2O_1C_3$	−122.0	−121.1
$H_7C_3C_4C_2$	−120.0	−121.5
$H_9C_4C_3C_2$	−90.0	−91.7

Distances in angstroms and angles in degrees.

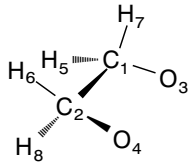
TABLE VI.
Behavior of Optimization Process for Triplet – Triplet Trans Conical Intersection for Carbon – Carbon Attack of Paterno – Buchi Reaction.

Iteration	RMS Gradient[a]	$\lvert E_{gs} - E_{ex}\rvert$[b]
1[c]	$1.74 \cdot 10^{-2}$	$7.01 \cdot 10^{-4}$
2[c]	$1.90 \cdot 10^{-2}$	$4.30 \cdot 10^{-4}$
3[c]	$1.58 \cdot 10^{-2}$	$3.10 \cdot 10^{-4}$
4[c]	$1.12 \cdot 10^{-2}$	$2.55 \cdot 10^{-4}$
5[c]	$5.20 \cdot 10^{-3}$	$2.23 \cdot 10^{-4}$
6	$5.14 \cdot 10^{-4}$	$1.43 \cdot 10^{-4}$
7	$6.70 \cdot 10^{-5}$	$1.28 \cdot 10^{-4}$
8	$4.20 \cdot 10^{-5}$	$1.28 \cdot 10^{-4}$
9	$3.40 \cdot 10^{-5}$	$1.12 \cdot 10^{-4}$
10	$2.50 \cdot 10^{-5}$	$9.60 \cdot 10^{-5}$
11	$2.50 \cdot 10^{-5}$	$9.60 \cdot 10^{-5}$
12	$2.50 \cdot 10^{-5}$	$9.60 \cdot 10^{-5}$
13	$1.70 \cdot 10^{-5}$	$8.00 \cdot 10^{-5}$
14	$1.70 \cdot 10^{-5}$	$8.00 \cdot 10^{-5}$
15	$1.70 \cdot 10^{-5}$	$8.00 \cdot 10^{-5}$
16	$1.70 \cdot 10^{-5}$	$8.00 \cdot 10^{-5}$
17	$1.70 \cdot 10^{-5}$	$6.4 \cdot 10^{-5}$
18	$8.00 \cdot 10^{-6}$	$6.4 \cdot 10^{-5}$
19	$8.00 \cdot 10^{-6}$	$6.4 \cdot 10^{-5}$

[a] RMS gradient in hartrees / bohr.
[b] Energy differences in hartrees.
[c] Iterations where the restricted step given by eqs. (21a, b) is active.

of the ground-state (S_0) 1,2-dioxetane to produce a biradical with a small activation energy; passage through an avoided crossing $S_0 - T_1$ in the oxygen–oxygen bond breaking just before the biradical minimum; passage through a second $S_0 - T_1$ real crossing just after the biradical minimum in the carbon–carbon bond breaking; passage through a transition state in the T_1 surface for carbon–carbon bond breaking to produce triplet (T_1) and ground-state (S_0) formaldehyde. We report here the geometry and the optimization process for the second real $S_0 - T_1$ crossing located in the carbon–carbon bond-breaking region. The search was carried out using a CI (4, 4)/AM1 wave function within the C_2 symmetry. The selected wave function correlates the more important valence electrons for the S_0 ground state (4π electrons) and the T_1 state (3π electrons). In Table VII the starting and final geometries are presented and Table VIII shows the behavior of the algorithm, which is the same as the previous examples. The initial geometry was that corresponding to the biradical minimum of the S_0 state. The BFGS formula was used for updating the Hessian matrix.

TABLE VII.
Geometry of Singlet – Triplet, $S_0 - T_1$, Intersection in Carbon – Carbon Bond-Breaking Region for 1,2-Dioxetane Decomposition within C_2 Symmetry.



Parameter	Initial	Final
C_1C_2	1.540	1.674
C_1O_3	1.282	1.293
H_5C_1	1.142	1.134
H_7C_1	1.144	1.120
$C_2C_1O_3$	114.7	111.8
$H_5C_1O_3$	113.5	111.6
$H_7C_1O_3$	113.3	115.6
$O_4C_2C_1O_3$	17.3	27.3
$H_5C_1O_3C_2$	118.3	110.8
$H_7C_1O_3C_2$	−117.7	−120.3

Distances in angstroms and angles in degrees.

TABLE VIII.
Behavior of Optimization Process of Singlet – Triplet, $S_0 - T_1$, Intersection in Carbon – Carbon Bond-Breaking Region for 1,2-Dioxetane Decomposition.

Iteration	RMS Gradient[a]	$\lvert E_{gs} - E_{ex}\rvert$[b]
1[c]	$4.19 \cdot 10^{-2}$	$3.35 \cdot 10^{-4}$
2[c]	$3.13 \cdot 10^{-2}$	$9.60 \cdot 10^{-5}$
3[c]	$1.93 \cdot 10^{-2}$	$3.20 \cdot 10^{-5}$
4[c]	$9.06 \cdot 10^{-3}$	0.00
5[c]	$7.22 \cdot 10^{-3}$	$1.60 \cdot 10^{-5}$
6[c]	$7.04 \cdot 10^{-3}$	$4.80 \cdot 10^{-5}$
7[c]	$5.88 \cdot 10^{-3}$	$1.59 \cdot 10^{-4}$
8[c]	$3.22 \cdot 10^{-3}$	$2.39 \cdot 10^{-4}$
9[c]	$1.86 \cdot 10^{-3}$	$2.55 \cdot 10^{-4}$
10[c]	$2.19 \cdot 10^{-3}$	$2.39 \cdot 10^{-4}$
11[c,d]	$2.30 \cdot 10^{-3}$	$2.23 \cdot 10^{-4}$
12[c]	$2.02 \cdot 10^{-3}$	$2.55 \cdot 10^{-4}$
13[c]	$1.95 \cdot 10^{-3}$	$2.23 \cdot 10^{-4}$
14[c,d]	$1.93 \cdot 10^{-3}$	$2.23 \cdot 10^{-4}$
15[c]	$1.80 \cdot 10^{-3}$	$1.91 \cdot 10^{-4}$
16[c]	$1.10 \cdot 10^{-3}$	$1.59 \cdot 10^{-4}$
17[c]	$4.81 \cdot 10^{-4}$	$1.59 \cdot 10^{-4}$
18[c]	$3.71 \cdot 10^{-4}$	$1.28 \cdot 10^{-4}$
19[c]	$2.78 \cdot 10^{-4}$	$1.28 \cdot 10^{-4}$
20[c]	$1.43 \cdot 10^{-4}$	$1.12 \cdot 10^{-4}$
21[c]	$1.35 \cdot 10^{-4}$	$1.12 \cdot 10^{-4}$
22[c]	$8.40 \cdot 10^{-5}$	$9.60 \cdot 10^{-5}$
23	$5.90 \cdot 10^{-5}$	$9.60 \cdot 10^{-5}$
24	$9.30 \cdot 10^{-5}$	$8.00 \cdot 10^{-5}$
25[c]	$9.30 \cdot 10^{-5}$	$8.00 \cdot 10^{-5}$
26	$5.90 \cdot 10^{-5}$	$6.40 \cdot 10^{-5}$
27	$1.70 \cdot 10^{-5}$	$6.40 \cdot 10^{-5}$

[a] RMS gradient in hartrees / bohr.
[b] Energy differences in hartrees.
[c] Iterations where the restricted step given by eqs. (21) is active.
[d] Iterations where the Powell device given by eq. (24) and (25) is active.

Note that the Powell device was active in iterations 11 and 14. The final converged gradient difference vector, $\partial(E_{gs} - E_{ex})/\partial \mathbf{q}$, corresponds to an increase of the C_1C_2 bond and a decrease of both CO bonds. At the final converged point the maximum gradient component in the S_{ti} subspace is $\max|(\mathbf{g}_{ti})_i| = 3.60 \cdot 10^{-5}$ hartrees/bohr. The geometry differs very little from that reported by Reguero et al.[28] except in the bond distance C_1C_2, which is about 0.1 Å bigger, and the CO bond distance, which is about 0.1 Å smaller.

EFFECT OF RESTRICTED STEP ON CONVERGENCE AND POWELL DEVICE

Except for example 4 the other examples of the Powell device, eqs. (24) and (25), are never active. This is because we are in the region where the restriction can be approximated very well in linear form with a positive curvature. Far from this point the curvature of the restriction can be negative, which affects the Hessian curvature of the Lagrangian function $\mathbf{W}_{ex}$. To illustrate this point and the effect of the restricted step on the final convergence, we repeated the optimization of the triplet–triplet trans conical intersection of the Paterno–Buchi reaction without the restricted step. The results are presented in Table IX. We observe

TABLE IX.
Behavior of Optimization Process for Triplet – Triplet Trans Conical Intersection for Carbon – Carbon Attack of Paterno – Buchi Reaction without Restricted Step Technique.

Iteration	RMS Gradient[a]	$\lvert E_{gs} - E_{ex}\rvert$[b]	θ[c]
1	$1.74 \cdot 10^{-2}$	$7.01 \cdot 10^{-4}$	1.00
2	$1.23 \cdot 10^{-2}$	$4.62 \cdot 10^{-4}$	1.00
3	$7.39 \cdot 10^{-3}$	$3.51 \cdot 10^{-4}$	1.00
4	$2.30 \cdot 10^{-2}$	$1.34 \cdot 10^{-3}$	1.00
5	$1.14 \cdot 10^{-2}$	$7.65 \cdot 10^{-4}$	1.00
6	$1.99 \cdot 10^{-2}$	$2.87 \cdot 10^{-3}$	0.61
7	$6.40 \cdot 10^{-2}$	$7.97 \cdot 10^{-4}$	1.00
8	$7.90 \cdot 10^{-2}$	$3.36 \cdot 10^{-3}$	0.29
9	$1.14 \cdot 10^{-1}$	$9.40 \cdot 10^{-4}$	1.00
10	$1.09 \cdot 10^{-1}$	$3.36 \cdot 10^{-3}$	0.43
11	$1.23 \cdot 10^{-1}$	$8.61 \cdot 10^{-4}$	1.00
12	$1.25 \cdot 10^{-1}$	$3.30 \cdot 10^{-3}$	0.48
13	$1.99 \cdot 10^{-1}$	$8.61 \cdot 10^{-4}$	1.00
14	$1.97 \cdot 10^{-1}$	$3.81 \cdot 10^{-3}$	0.26
15	$2.23 \cdot 10^{-1}$	$9.72 \cdot 10^{-4}$	1.00
16	$1.26 \cdot 10^{-1}$	$2.95 \cdot 10^{-3}$	1.00
17	$1.64 \cdot 10^{-1}$	$7.81 \cdot 10^{-4}$	1.00
18	$2.91 \cdot 10^{-2}$	$1.47 \cdot 10^{-3}$	1.00
19	$2.77 \cdot 10^{-3}$	$1.28 \cdot 10^{-4}$	1.00
20	$1.00 \cdot 10^{-3}$	$1.12 \cdot 10^{-4}$	1.00
21	$5.90 \cdot 10^{-4}$	$9.60 \cdot 10^{-5}$	1.00
22	$4.97 \cdot 10^{-4}$	$9.60 \cdot 10^{-5}$	1.00
23	$4.89 \cdot 10^{-4}$	$9.60 \cdot 10^{-5}$	1.00
24	$3.88 \cdot 10^{-4}$	$8.00 \cdot 10^{-5}$	1.00
25	$2.70 \cdot 10^{-4}$	$8.00 \cdot 10^{-5}$	1.00
26	$1.35 \cdot 10^{-4}$	$8.00 \cdot 10^{-5}$	1.00
27	$5.10 \cdot 10^{-5}$	$8.00 \cdot 10^{-5}$	1.00
28	$1.70 \cdot 10^{-5}$	$6.40 \cdot 10^{-5}$	1.00

[a] RMS gradient in hartrees / bohr.
[b] Energy differences in hartrees.
[c] Value defined in eq. (25).

that the effect of the restricted step is crucial to reach the convergence, otherwise one gets a poor convergence. We note the oscillatory behavior of the RMS gradient. On the other hand, we see that the Powell device is active in iterations 6, 8, 10, 12, and 14 (the parameter θ is different from 1). Generally, the Powell device is active in the iterations where the values of the restriction, $|E_{gs} - E_{ex}|$, increase.

USE OF MSP FORMULA FOR UPDATING HESSIAN MATRIX OF LAGRANGIAN FUNCTION

In Table X we present a comparison of the number of iterations employed by the algorithm using the BFGS[12] or the MSP[13b,21] formulae for each example presented above [see eqs. (18) and (19), respectively]. Clearly when the MSP update formula is used, the algorithm needs a greater number of iterations. This result is easily justified because we are minimizing the Lagrangian function given in eq. (6) and it is well known that for minimization the BFGS formula is much better that the MSP formula.[12,21]

Summary and Conclusions

Within the general problem of finding an efficient algorithm for locating a minimum energy crossing point between two potential energy surfaces, this article concerns two basic aspects. First, we present an optimization of a Lagrangian function coupled with the use of the restricted step applied in the subspace of the independent variables, namely the S_{ti} subspace. In this way, when the algorithm fulfills the constraints but is still far from the minimum point, the restricted step insures to some degree that the minimization is carried out in the feasibility region. Also, it warrants that the Hessian in the S_{ti} subspace is positive definite, insuring the search for a minimum. Second, we present a short numerical study on the best updated Hessian formula for Lagrangian optimization. Because we are concerned with a minimum of the Lagrangian function, the BFGS formula works very well, taking into account the Powell device. The examples presented show that the method is capable of locating a minimum energy crossing point between surfaces of different electronic states. The proposed algorithm can be easily generalized for locating m-fold degenerate minimum energy crossing points.

TABLE X.
Comparison Between Total Number of Iterations Employed by Algorithm Using BFGS and MSP Formulae for Updating of Hessian Matrix W_{ex}.

	Example 1	Example 2	Example 3	Example 4
BFGS[a]	19	6	19	28
MSP[b]	43	11	20	34

[a] Hessian matrix updated according to eq. (18).
[b] Hessian matrix updated according to eq. (19).

Acknowledgment

We are indebted to Professor S. Olivella for his valuable suggestions. This research was supported by the Spanish DGICYT (Grant PB92-0796-C01-02).

References

1. J. von Neumann and E. Wigner, *Phys. Z.*, **30,** 467 (1929).
2. (a) E. Teller, *J. Phys. Chem.*, **41,** 109 (1937); (b) G. Herzberg and H. C. Longuet–Higgins, *Trans. Faraday Soc.*, **35,** 77 (1963); (c) J. C. Tully and R. K. Preston, *J. Am. Chem. Soc.*, **55,** 562 (1971); (d) L. Salem, *Electrons in Chemical Reactions: First Principles*, Wiley, New York, 1987.
3. N. Koga and K. Morokuma, *Chem. Phys. Lett.*, **119,** 371 (1985).
4. P. J. Kuntz and W. N. Whitton, *J. Chem. Phys.*, **95,** 5149 (1991).
5. A. Farazdel and M. Dupuis, *J. Comput. Chem.*, **12,** 276 (1991).
6. I. N. Ragazos, M. A. Robb, F. Bernardi, and M. Olivucci, *Chem. Phys. Lett.*, **197,** 217 (1992).
7. M. R. Mana and D. R. Yarkony, *J. Chem. Phys.*, **99,** 5251 (1993).
8. D. R. Yarkony, *J. Chem. Phys.*, **97,** 4407 (1993).
9. M. J. Bearpark, M. A. Robb, and H. B. Schlegel, *Chem. Phys. Lett.*, **223,** 269 (1994).
10. P. Celani, M. A. Robb, M. Garavelli, F. Bernardi, and M. Olivucci, *Chem. Phys. Lett.*, **243,** 1 (1995).
11. M. Klessinger, *Angew. Chem. Int. Ed. Engl.*, **34,** 549 (1995) and references therein.
12. R. Fletcher, *Practical Methods Of Optimization*, 2nd ed., Wiley, New York, 1987.
13. (a) P. Culot, G. Dive, V. H. Nguyen, and J. M. Ghuysen, *Theor. Chim. Acta*, **82,** 189 (1992); (b) J. M. Bofill, *J. Comput. Chem.*, **15,** 1 (1994).
14. C. A. Mead and D. G. Truhlar, *J. Chem. Phys.*, **70,** 2284 (1979).

15. G. J. Atchity, S. S. Xantheas, and K. Ruedenberg *J. Chem. Phys.*, **95,** 1862 (1991).
16. J. Katriel and E. R. Davidson, *Chem. Phys. Lett.*, **76,** 259 (1980).
17. S. P. Han, *Math. Prog.*, **11,** 263 (1976).
18. M. J. D. Powell, *Math. Prog.*, **14,** 224 (1978).
19. D. Gabay, In *Mathematical Programming Study 16*, A. G. Buckley and J.-L. Goffin, Eds., North-Holland Publishing Company, Amsterdam, 1982.
20. P. E. Gill and W. Murray, In *Numerical Methods for Constrained Optimization*, P. E. Gill and W. Murray, Eds., Academic Press, London, 1974.
21. J. M. Bofill and M. Comajuan, *J. Comput. Chem.*, **16,** 1326 (1995).
22. AMPAC program, local version, extended by D. Liotard, 1987.
23. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, *J. Am. Chem. Soc.*, **107,** 3902 (1985).
24. M. J. S. Dewar and D. Liotard, *J. Mol. Struct. (Theochem.)*, **206,** 123 (1990).
25. I. J. Palmer, I. N. Ragazos, F. Bernardi, M. Olivucci, and M. A. Robb, *J. Am. Chem. Soc.*, **116,** 2121 (1994).
26. B. Bigot, R. Ponec, A. Sevin, and A. Devaquet, *J. Am. Chem. Soc.*, **100,** 6575 (1978).
27. N. Yamamoto, F. Bernardi, A. Bottoni, M. Olivucci, M. A. Robb, and S. Wilsey, *J. Am. Chem. Soc.*, **116,** 2064 (1994).
28. M. Reguero, F. Bernardi, A. Bottoni, M. Olivucci, and M. A. Robb, *J. Am. Chem. Soc.*, **113,** 1566 (1991).